原 浙大《概率与梳理统计》第四版证明随机变量X,Y的相关系数的绝对值小于1，及一些疑问

这里的证明方法来自浙大《概率与梳理统计》108页：
需证明命题：
对任意两个随机变量X, Y，证明其相关系数的绝对值小于1。

证明思路：

1. 先构造$a+bX$和Y的均方误差的期望$E\left\{[Y-(a+bX)]^2\right\}$
2. 再求出这个期望的最小值是$\left(1-\rho_{XY}^2\right)D(Y)$
3. 因为均方误差是大于等于0的值，方差也是大于等于0的值，所以对于等式

$$\min_{a,b} E\left\{[Y-(a+bX)]^2\right\}=\left(1-\rho_{xy}^2\right)D(Y)$$

$\left(1-\rho_{XY}^2\right)$是必须大于0的，所以$|\rho_{xy}|\leq 1$

具体步骤
1，对于任意两个随机变量X, Y，构造

$$\begin{aligned} e &= E\left\{[Y-(a+bx)]^2\right\} \\ &= E\left(Y^2\right)+b^2E\left(X^2\right)+a^2-2bE(XY)+2abE(X)-2aE(Y) \end{aligned} \tag{1}$$

2. 可以将e看作是关于a, b的函数，那么，根据多元函数求极值的方法，就要分别求e关于a, b的偏导，并令其等于0, 找到满足必要条件的a, b的值

$$\begin{cases} \dfrac{\partial e}{\partial a}=2a+2bE(X)-2E(Y)=0, \\ \dfrac{\partial e}{\partial b}=2bE\left(X^2\right)-2E(XY)+2aE(X)=0. \end{cases} \tag{2}$$

求a, b的二元一次方程组，得出

$$\begin{aligned} b_0 &= \frac{Cov(X,Y)}{D(X)} \\ a_0 &= E(Y)-b_0E(X)=E(X)-E(X)\frac{Cov(X,Y)}{D(X)} \end{aligned} \tag{3}$$

3. 将a_0, b_0代入式(1)中，得到

$$\begin{aligned} e &= E\left\{[Y-(a+bx)]^2\right\} \\ &= D[y-a_0-b_0X]+[E(Y-a_0-b_0X)]^2 \\ \text{由(2)的第一式可得，} E(Y-a_0-b_0X)=0 \\ &= D(Y-b_0X) \\ &= D(Y)+b_0^2D(X)-2b_0Cov(X,Y) \\ &= D(Y)+\frac{Cov^2(X,Y)}{D(x)}-2\frac{Cov^2(X,Y)}{D(x)} \\ &= D(Y)\left[1-\frac{D(X,Y)}{D(X)D(Y)}\right] \\ &= \left(1-\rho_{XY}^2\right)D(Y) \end{aligned}$$

因为由方差的定义可知，方差始终是大于或等于0的值，而且e是一个非负随机变量(有平方)的期望，所以$\left(1-\rho_{XY}^2\right)$必定大于0，所以$X, Y$的相关系数$|\rho_{XY}|\leq 1$

这里的证明过程没有问题，不过我有疑问的是，如果问题改为求任意随机变量X,Y的相关系数的取值范围，又该怎么做呢？因为这里的题目已经预设了证明是绝对值小于1，那么怎么证明1就是任意两个随机变量的相关系数的最大值呢？就像a是一个$(0, 0.5)$之间的值，现在已经证明了$a\leq 1$，怎么能够证明出$a\leq 0.5$呢

更新，一些关于相关系数的理解：

1. 相关系数表达的是两个随机变量X, Y的线性关系，不相关意味着两随机变量间没有线性关系，而不表示两随机变量独立,即是说不存在$a+bX=Y$这种关系，可能存在$Y=X^2$时，相关系数$\rho_{XY}=0, X, Y$不相关，但是X, Y也不独立。详见浙大概率论与梳理统计第4版108页及其后例一
2. 疑问解答：
   怎么证明出任意随机变量的相关系数的最大值就是1的呢？
   两个思路：

- 由知乎大佬的一个回答知道，可以将随机变量看作无穷维的向量，协方差看作是向量内积，那么相关系数就是向量的夹角余弦值，自然而然，取值范围是$[-1, 1]$
- 由书上108页的第二个定理的证明可以知道，$|\rho_{XY}|=1$的充要条件是存在常数a, b使得

$$P\{Y=a+bX\}=1$$

即是当且仅当$Y=a+bX$这个条件存在时，满足$|\rho_{XY}|=1$这个条件，对于其他$Y\neq a+bX$的情况都是$|\rho_{XY}|<1$，所以能够得出对于任意两个随机变量的相关系数，其绝对值最大能够取到1，以及$|\rho_{XY}|\leq 1$